Exploring sonic interaction design and presence: Natural Interactive Walking in Porto

Nordahl, Rolf; Serafin, Stefania; Fontana, Frederico

Published in:
Proceedings of Presence 2009

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Exploring sonic interaction design and presence: Natural Interactive Walking in Porto.

Rolf Nordahl, Stefania Serafin
Medialogy, Aalborg University Copenhagen
Lautrupvang 15, 2750 Ballerup, DK
rn,sts@media.aau.dk

Federico Fontana
Dipartimento di Informatica
University of Verona
fontana@sci.univr.it

ABSTRACT
In this paper we report on the results of a three days workshop whose goal was to combine interactive sounds and soundscape design to simulate the sensation of walking in a specific location of Porto.

We discuss advantages and disadvantages of the different solutions proposed in terms of the technology used, and issues of how sonic interaction combined with soundscape design affects presence in virtual environments.

Author Keywords
Sonic interaction, soundscape design, sounds of Porto.

INTRODUCTION
One of the most interesting ways of exploring a city is by walking. From the sonic point of view, walking also allows to express the natural landmark of a place.

When exploring a place by walking, two main categories of sounds can be identified: the person's own footsteps and the surrounding soundscape. Footsteps sounds represent important elements in movies and computer games. In these media, sounds are usually acquired from libraries or recorded by so-called Foley artists that put shoes in their hands and interact with different materials to simulate the act of walking. Recently, several algorithms have been proposed to simulate the sounds of walking algorithmically. One of the pioneers in this field is Perry Cook, who proposed a collection of physically informed stochastic models (PhiSM) simulating several everyday sonic events [3]. Among such algorithms the sounds of people walking on different surfaces were simulated [4]. A similar algorithm was also proposed in [6], where physically informed models simulate several stochastic surfaces.

Recently, in [5] a solution based on granular synthesis was proposed. The characteristic events of a footstep sounds were reproduced by simulating the so-called ground reaction force,i.e., the reaction force supplied by the ground at every step.

Studies on soundscape originated with the work of R. Murray Schafer [11]. Among other ideas, Schafer proposed soundwalks as empirical methods for identifying a soundscape for a specific location. In a soundwalk people are supposed to move in a specific location, noticing all the environmental sounds heard. Schafer claimed that each place has a soundmark,i.e., sounds which one identifies a place with. The idea of experiencing a place by listening has been recently further developed by Blesser and Salter [1]. By synthesizing technical, aesthetical and humanistic considerations, the authors describe the field of aural architecture and its importance in everyday life.

In the field of virtual reality, studies have recently shown how the addition of auditory cues could lead to measurable enhancement in the feeling of presence. Results are available on sound delivery methods [12, 10] or sound quality [2, 10]. Recently, the role of self-sound to enhance sense of presence in virtual environments has been investigated. By combining different kinds of auditory feedback consisting of interactive footsteps sounds created by ego-motion with static soundscapes, it was shown how motion in virtual reality is significantly enhanced when moving sound sources and ego-motion are rendered [7, 8].

The results presented in this paper are part of the Natural Interactive Walking (NIW) FET-Open project[1], whose goal is to provide closed-loop interaction paradigms enabling the transfer of skills that have been previously learned in everyday tasks associated to walking. In the NIW project, several walking scenarios are simulated in a multimodal context, where especially audition and haptics play an important role.

As part of the training sessions of the 2009 Sound and Music Computing Summer school, a workshop was organized. The workshop followed the strategy of combining interactive sounding objects [9] with soundscape design and the role of sound to create a sense of place was addressed. The main goal of the workshop was to allow students to experience the technology developed and to integrate it in the context of the city of Porto.

Presence 2009, Nov 11 – 13, 2009, Los Angeles, California, USA.

[1] http://www.niwproject.eu/

THE WORKSHOP

The workshop took place in Porto, Portugal, between July 18th and 21st. It lasted four afternoons, and 20 students participated. The first afternoon of the school was dedicated to forming groups. At the end of the afternoon, three students chose the natural interactive walking in Porto workshop. Students were assigned the task to record different soundscapes in Porto, and to combine them with interactive simulations of footsteps sounds as described in the following section. The ultimate goal of the workshop was the recreation in a laboratory setting of the sensation of walking in a specific location of Porto. During the second afternoon the soundscapes of Porto were recorded, as described later. In the third afternoon such soundscapes were combined with the interactive footsteps devices described in the following section. The last afternoon was dedicated to final tuning of the system and demonstrations to the other students.

SONIC INTERACTION DESIGN

When designing an augmented walking surface the physical configuration of the device, the nature of the control system, the sensors and actuators involved, as well as the rendering algorithm need to be taken into account. For example, two basic physical configurations of an augmented walking device can be envisaged (Fig. 1). In our situation, the first configuration consisted of shoes instrumented with sensors attached to them. The second configuration consisted of a rigid surface instrumented with microphones.

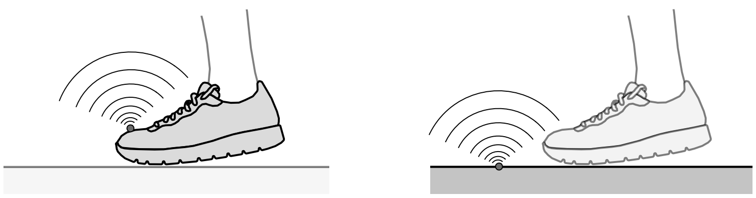

Figure 1. A walking surface augmented using an instrumented shoe (left) or with an instrumented floor (right).

THE FIRST SETUP: SHOES AUGMENTED WITH SENSORS

The first proposed prototype is shown in Figure 2. It provides a technological platform for the development of a multimodal wearable shoe-based interface. The hardware system consists of three main components, respectively providing: sensing, force data acquisition and conditioning, and sound computation and display.

The audio display is provided by two 4Ω Visaton FRWS5 loudspeakers, which were chosen for their low weight and small dimensions. Each speaker is attached onto a shoe, on the surface of the instep. By employing the laptop battery to power the acquisition and computation systems, plus a 9V battery to supply the small loudspeakers, the problem of wearability is addressed. On the other hand, in order to make the system wearable, the loudspeakers have to be small and light, this way raising problems of poor sound quality especially at low frequencies. This problem is addressed by enhancing the system with a portable subwoofer.

Figure 2. A pair of sabots with sensors in the soles. The loudspeakers are fixed to the shoes instep.

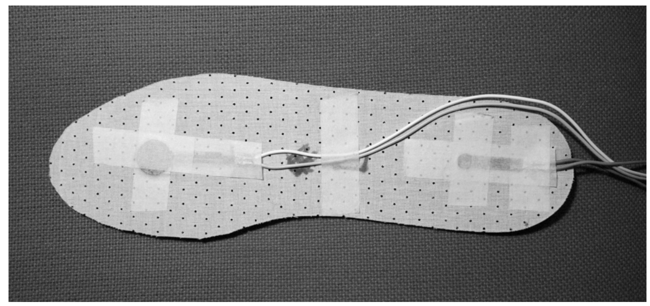

Figure 3. An insole with sensors. In this case, the loudspeaker can be fixed to the ankle with some scotch tape.

Force sensing

Two force sensing resistors (FSRs) have been used in each shoe sole: one for the toe (Interlink[2] 402 FSR) and one for the heel (Interlink 400 FSR). Although these two devices result too sensitive in detecting the weight of a human body, this issue is not a major concern as the sound processing system is enabled by force variations across time rather than absolute force values. The Arduino Duemilanove[3] USB board has been chosen for acquisition purposes.

Usability issues suggest to follow different approaches to make users in condition to wear sensing shoes. For this reason, we have inserted FSRs inside the soles (see Figure 2), as well as put them in the lower side of removable insoles (see Figure 3). The latter solution, in particular, allows to reuse the same sensors in different shoes.

Sound processing and auditory display

The synthesis engine runs on a laptop which computes sounds in real-time according to the algorithm described in Section . The signal so synthesized is sent to a TDA2822 audio power amplifier chip which is supplied by a $9V$ / $15mAh$ battery. Both the laptop and the amplifier are placed in a backpack.

Software for sound synthesis

The processing software conditions the sensed force input, and synthesizes sounds. To make it run a on a general-purpose architecture, the software has been written as a com-

[2] `www.interlinkelectronics.com`
[3] `www.arduino.cc`

bination of C externals and programs (patches) for the multi-platform real-time DSP environment Pure Data[4] (Pd).

To date, a physically-based crumpling sound model has been tuned to simulate aggregate grounds such as thick snow, gravel, creaking wood.

The model consists of a control layer and a sound synthesis layer implemented as a Pd patch embedding the C externals. More specifically, crumpling sounds are obtained as a result of a stochastic temporal process (control layer) which drives an impact model (sound synthesis), namely triggering impacts with different velocity and at different instants in time. The control layer takes two parameters as input: *force*, which is directly linked to the force data and sets the instantaneous power that enters the overall system, and *resistance*, which determines how resisting the simulated floor materials are, on average, against the incoming force before their state changes. Concerning the physical attributes of the model, parameters of *modal frequencies*, *decay times*, *impact stiffness*, and *shape of the contacting surface* are exposed to control.

Four instances of the model – one for each FSR – are included in separate Pd sub-patches which receive the filtered force data.

THE SECOND SETUP: FLOOR MICROPHONES

We adopted a set of non-contact microphones placed on the floor. In our experiments we used the Shure BETA 91[5], a high performance condenser microphone with a tailored frequency response designed specifically for kick drums and other bass instruments. Its features made it a good candidate for our purpose of capturing the footsteps sounds. In our experiments we placed two microphones on the floor at 1.5 meters distance from each other.

Extraction of the ground reaction force

A footstep sound is the result of multiple micro-impact sounds between the shoe and the floor. The set of such micro-events can be thought as an high level model of impact between an *exciter* (the shoe) and a *resonator* (the floor). In such a vision the sound captured by the microphones can be considered as a composition of both these two components.

In mechanics such exciter is usually called ground reaction force (GRF), i.e., the reaction force supplied by the ground at every step. The aim of the phase of analysis has been that of finding some parameters that allowed us to extrapolate the exciter from the captured sound, i.e., finding the ground reaction force from the acoustic waveform. Such an extrapolation consisted in removing, from the spectral representation of the sound, the main modes.

In order to achieve the final goal of producing the impression of walking on floors made of different kinds of materials, we removed the contribution of the resonator, kept the exciter and considered the latter as input for a new resonator that implements different kinds of floors. Subsequently the contribution of the shoe and of the new floor are summed in order to have a complete footstep sound.

The algorithm has been implemented in real-time as an extension to the Pd platform and works as follows: the sound of a person walking is detected in real-time by the microphones described above. From this sound, the resonances corresponding to the impact of the shoe on the floor are removed, in order to extract the GRF. Such GRF is used as input to the sound synthesis algorithms described in the following section.

Sound synthesis and manipulations

The GRF estimated with the technique described in the previous section was used to control two different sound synthesis algorithms, reproducing solid and aggregate surfaces respectively.

To synthesize solid surfaces, we used an algorithm which physically simulates in real-time the contact between hard surfaces. In particular, we controlled one of the input parameters of the model, the impact force, by using the estimated GRF as described in the previous section. This allowed us to recreate realistic footsteps sounds. By varying the different parameters of the model it was possible to simulate the sounds of different surfaces, although a systematic mapping of physical to perceptual parameters is not in place yet.

To synthesize aggregate surfaces, we implemented the physically informed sonic models (PhiSM) [3]. The stochastic energy of the models is controlled by using the estimated GRF.

SOUNDSCAPE RECORDINGS

During the second day of the workshop, the group of students and tutors visited different locations in Porto to record characteristic soundscapes. The tour started from a shopping mall in the outskirt of the city, to the little streets in the central area, one of the main bridges, the metro and the skateboarders playing in front of the Casa da Musica building. All recordings were performed using Zoom H4 or H2 recorders [6]. The students were instructed to record at least three characteristic locations in Porto, and to bring along the recordings the following day.

The third afternoon started with listening the different recordings and discussing why they were interesting and representative of the sounds of Porto. We then encourage the students to produce an overall soundscape to be combined with the interactive footsteps sounds previously described.

PUTTING IT ALL TOGETHER

The integration of the interactive footsteps sounds with the final soundscape was straightforward from the technical point of view, thanks to the capabilities offered by the Pure Data software. In each section of the soundscape, the interactive footsteps sounds were chosen in such a way to match the

[4] www.puredata.info

[5] http://www.shure.com/

[6] www.zoom.co.jp

particular location. As an example, when the soundscape was recorded at the beach, footsteps sounds of walking on sands were triggered. Due to the limited amount of time, the spatial characteristic of the place were not fine tuned to the corresponding footsteps. In other words, time limitations did not allow to add suitable reverberation effects to the interactive footsteps sounds.

ADVANTAGES AND DISADVANTAGES OF THE TWO SETUPS

Among other things, the workshop provided the possibility to compare the two developed systems and understand their advantages and disadvantages.

Portabilty

Both systems are easily portable. The first system consists of a pair of shoes enhanced with sensors, a subwoofer and a laptop. The second system consists of two microphones, a soundcard and a laptop.

Easiness of setup

To setup the first system it is necessary to ensure that the Arduino is receiving data from the sensors and sending them to the Pure Data environment. Once this step is performed, after a bit of calibration of the sensors data the system is ready to run. The second setup requires the microphone to send data to the Fireface audio interface which sends it to the Pure data environment. So in both situations the easiness of setup stands on the reliability of the interfaces.

Sound quality

The sound quality of the systems obviously depends on the quality of the sound synthesis algorithms as well as the audio delivery methods used. Since the whole setup is intended to reproduce ecological walking sounds resulting from the mechanical interaction between shoes and ground, high sound pressure is not required. Conversely, exceedingly poor low frequencies make the system unsuitable for the rendering of audible low resonances as those elicited when a foot bumps over a hollow floor like a cavity covered with wooden bars. The second setup is delivered through headphones. This is due to the fact that the surrounding sonic environment needs to be relatively quiet, since the microphones should pick up only the real footsteps sounds.

As concerns the quality of the synthesized sounds using the microphone based system, good results have been obtained from the synthesis of big and little gravel, as well as different kinds of wood. In both cases, further improvements to the sound synthesis algorithms are under development.

Sensing capabilities

The shoe enhanced interface forms a solid base where to build further shoe-based functionalities. The reliability of the sensing and conditioning stage, along with the low-latency signal processing and sufficiently realistic auditory display, have already resulted into a credible closed-loop interaction. The speakers and the battery-powered amplifier together give rise to a noticeable lack of low frequencies, neither they are able to handle sharp transients nor large dynamic ranges. This problem has been addressed by the use of the additional subwoofer. The use of the microphones on the ground did not show any particular problem, and the GRF was extracted in a clear way. The limitation of this system is that only the audio waveform can be obtained, and from that some information extracted in real-time, such as the GRF. As an example, the velocity information can be hardly detected from the acoustic waveform.

However, the envelope extraction as GRF turned out to be the right choice for the control of some parameters of the sound synthesis algorithms.

Wearability

The shoe enhanced system requires users to wear a specific size of footwear, and require a laptop to be carried on the back. On the other end, the floor microphones allow the users to keep their own footwear.

Navigation

The shoe enhanced system ideally allow to navigate without restriction, since the all system is integrated inside the laptop which the users wear. On the other end, the floor microphones require the user to navigate in a specific location delimited by the space inside the microphones.

Integration in VR environments

Both systems have been developed as extension to the Pure Data platform. The platform is open source and can be easily combined with several interfaces and different software packages. A protocol which has been shown to be suitable for integration purposes is the Open Sound Control protocol [7]. In the future, both systems will be integrated with haptic and visual feedback, to simulate different multimodal environments.

CONCLUSIONS AND FUTURE WORK

In conclusion, both systems showed to be suitable as a floor based interaction device to navigate virtual environments. The shoes enhanced with sensors have the advantage of a higher number of sensing capabilities and do not require the environment to be acoustically isolated. On the other end, they require users to be able to wear a particular size of shoe and to wear a backpack with a laptop. The floor microphones have the advantage that users can wear their own shoes. However, they require a quiet environment to be used, and they have limited sensing capabilities.

The integration of the systems with soundscapes of Porto proved to be successful. Although not formally tested, the participants to the workshop enjoyed to be able to virtually walk around the city listening to some characteristic sonic landmarks. Moreover, the students enrolled in the workshop found it valuable to provide a context to the interactive footsteps to create a more imersive experience. The first prototypes of the complete setups could be easily prepared even

[7] opensoundcontrol.org

given the short duration of the workshop. However, their integration can be significantly improved, especially by adding a sense of space to the rendering of the interactive footsteps sounds, and by adding interactivity also to the soundscapes of the city, which right now was simply a static recording.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Communitys Seventh Framework Programme under FET-Open grant agreement n 222107 NIW - Natural Interactive Walking and the COST Action Sonic Interaction Design (SID) (IC0601).

REFERENCES

1. B. Blesser and L. Salter. *Spaces speak, are you listening?: experiencing aural architecture*. MIT Press, 2006.

2. P. Chueng and P. Marsden. Designing Auditory Spaces to Support Sense of Place: The Role of Expectation. In *CSCW Workshop: The Role of Place in Shaping Virtual Community*. Citeseer, 2002.

3. P. Cook. Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, 21(3):38–49, 1997.

4. P. Cook. Modeling Bills Gait: Analysis and Parametric Synthesis of Walking Sounds. *Proceedings of the AES 22nd International Conference on Virtual, Synthetic, and Entertainment Audio*, pages 73–78, 2002.

5. A. Farnell. Marching onwards: procedural synthetic footsteps for video games and animation. . *Proceedings of the Pure Data Convention*, 2007.

6. F. Fontana and R. Bresin. Physics-based sound synthesis and control: crushing, walking and running by crumpling sounds. In *Proc. Colloquium on Musical Informatics*, pages 109–114, Florence, Italy, May 2003.

7. R. Nordahl. Self-induced Footsteps Sounds in Virtual Reality: Latency, Recognition, Quality and Presence. In *Presence*, 2005.

8. R. Nordahl. Increasing the Motion of Users in Photo-realistic Virtual Environments by Utilising Auditory Rendering of the Environment and Ego-motion. In *Presence*, 2006.

9. D. Rocchesso, R. Bresin, and M. Fernstróm. Sounding objects. *IEEE MULTIMEDIA*, pages 42–52, 2003.

10. R. Sanders Jr. *The effect of sound delivery methods on a users sense of presence in a virtual environment*. PhD thesis, NAVAL POSTGRADUATE SCHOOL, 2002.

11. R. Schafer. *The tuning of the world*. Random House Inc, 1977.

12. R. Storms and M. Zyda. Interactions in perceived quality of auditory-visual displays. *Presence: Teleoperators & Virtual Environments*, 9(6):557–580, 2000.